Hearing Research xxx (2014) 1–10

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Deep electrode insertion and sound coding in cochlear implants

Ingeborg Hochmair [a,*], Erwin Hochmair [b], Peter Nopp [a], Melissa Waller [a], Claude Jolly [a]

[a] MED-EL GmbH, Fürstenweg 77a, A-6020 Innsbruck, Austria
[b] Institute for Ion Physics and Applied Physics, University of Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria

ARTICLE INFO

Article history:
Received 3 July 2014
Received in revised form
8 October 2014
Accepted 10 October 2014
Available online xxx

ABSTRACT

Present-day cochlear implants demonstrate remarkable speech understanding performance despite the use of non-optimized coding strategies concerning the transmission of tonal information. Most systems rely on place pitch information despite possibly large deviations from correct tonotopic placement of stimulation sites. Low frequency information is limited as well because of the constant pulse rate stimulation generally used and, being even more restrictive, of the limited insertion depth of the electrodes. This results in a compromised perception of music and tonal languages.

Newly available flexible long straight electrodes permit deep insertion reaching the apical region with little or no insertion trauma. This article discusses the potential benefits of deep insertion which are obtained using pitch-locked temporal stimulation patterns. Besides the access to low frequency information, further advantages of deeply inserted long electrodes are the possibility to better approximate the correct tonotopic location of contacts, the coverage of a wider range of cochlear locations, and the somewhat reduced channel interaction due to the wider contact separation for a given number of channels.

A newly developed set of strategies has been shown to improve speech understanding in noise and to enhance sound quality by providing a more "natural" impression, which especially becomes obvious when listening to music.

The benefits of deep insertion should not, however, be compromised by structural damage during insertion. The small cross section and the high flexibility of the new electrodes can help to ensure less traumatic insertions as demonstrated by patients' hearing preservation rate.

This article is part of a Special Issue entitled <Lasker Award>.

1. Introduction

A sound signal can be divided into two principal components, referred to as the envelope and the (temporal) fine structure (Hilbert, 1912). The envelope of a signal is defined by the Hilbert transform. It can be approximated by rectification followed by a low pass filter. The fine structure contains information about instantaneous frequency of a sound and is coded in the time domain via phase locking in the low frequencies. In speech and other acoustic signals, envelope and fine structure contribute differentially to the comprehension of sounds. Smith et al. (2002) and later Xu and Pfingst (2003) quantified these relative contributions as a function of the number of analyzed filter bands (corresponding to the number of channels in a cochlear implant). Results revealed that speech perception largely relies on the envelope of the sound whereas music and other tonal instances of sounds like prosody or tonal languages are mainly conveyed by the fine structure of the sound signal. This already hints at the improvements to be expected from apical temporal coding: "naturalness", better performance with tonal languages, and more enjoyable perception of music. The degree of "naturalness" can be described by a single sided deaf subject by comparing the electrically versus the acoustically generated impression.

It is well known from physiology that – depending on frequency – sounds are not only coded in cochlear place but also in the temporal structure of neural responses, referred to as the time code. In natural hearing, low frequency sounds are coded both in place and time in the apical region of the cochlea. Sound frequency is thus not only coded in place but is also reflected in the temporal

Abbreviations: EAS, electric-acoustic stimulation; CI, cochlear implant; CIS, continuous interleaved sampling; CIS+, CIS-variant; FSP, FS4, FS4-p, fine structure strategies; HDCIS, CIS-variant; HP, hearing preservation; SRT, speech reception threshold; MCL, most comfortable loudness; VEMP, vestibular evoked myogenic potential

* Corresponding author. Tel.: +43 664607055100.
E-mail address: ingeborg.hochmair@medel.com (I. Hochmair).

http://dx.doi.org/10.1016/j.heares.2014.10.006
0378-5955/© 2014 Published by Elsevier B.V.

neural response pattern here. With increasing frequency, time coding vanishes, so that high-frequency sounds are only coded in place – the temporal response pattern of the neurons no longer reflects sound frequency. Temporal coding is produced by a mechanism that is usually referred to as phase locking, meaning that neural responses tend to arise at a certain point in time during each single period of the stimulus. With increasing frequency, phase locking and thus time coding vanish at frequencies beyond approximately 1.5 kHz.

In normal hearing, time coding and place coding usually covary in the cochlea so that it has been difficult to assess the importance for the low frequencies of either of these codes in isolation. Research using transposed tones (Oxenham et al., 2004), however, has demonstrated that when low-frequency sinusoids are presented to places in the cochlea that are tuned to higher frequencies, i.e. in the case of a mismatch between time code and place code, then pitch perception deteriorates dramatically when compared to the matched-time-place condition. In addition, the ability to extract the pitch (i.e. fundamental frequency) of a sound from a multitude of low-frequency harmonics disappears if these low-frequency harmonics are presented to high-frequency places in the cochlea. All in all, these results demonstrate the importance of frequency-place matching. Consequently, with electrical stimulation the mapping of frequency bands to location influences the "naturalness" of the elicited sensations.

2. Coding strategies

Strategies and algorithms for representing sounds through a cochlear implant have been a core challenge in cochlear implants from the early days. In the early '80's the more fundamental questions, like.

monopolar or bipolar stimulation.
analog or pulsatile stimulation.
whole signal presentation or feature extraction.
fine temporal structure or place pitch.
had to be addressed.

Our first design was a multichannel implant intended for pulsatile stimulation (I. Hochmair, 2013). Having been implanted in

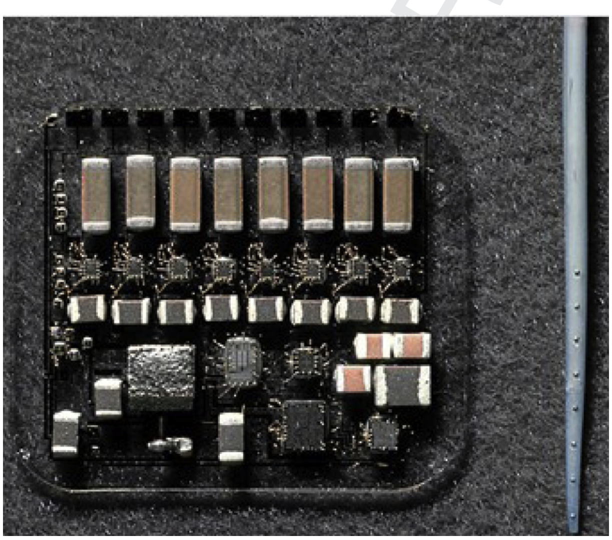

Fig. 1. 8-channel microelectronic cochlear implant together with scala tympani electrode. The substrate containing the electronic components was encased within a hermetic glass package and connected to the electrode and to the receiver coil (not shown here) via hermetic feed throughs.

Dec. 1977, it was the first microelectronic multichannel cochlear implant (Fig. 1). It thus may be considered as the prototype of the modern cochlear implant. However, our experience with it led us to the conclusion that we needed a more signal-transparent system that would give us more flexibility in developing a viable coding strategy. To avoid percutaneous plugs, which are great for research but rather burdensome for the patient, we developed a passive transcutaneous four channel system which turned out to be our workhorse for the coming years. Since it allowed stimulation with any kind of pulsatile or analog waveform, it opened the door to a whole new realm of research possibilities. It was extensively used in our laboratory work to run psychoacoustic tests as well as to explore the possibilities of multichannel coding strategies. For the wearable processor only one channel was used. It took us almost 12 more years to reassume our original approach.

To keep power consumption low, we had quickly decided to use monopolar stimulation, despite findings from animal experiments demonstrating the narrower stimulation range of bipolar stimulation. This decision more or less answered the remaining questions: a large current spread around intrascalar stimulation contacts is less amenable to multichannel stimulation providing place pitch, but rather to a single channel broadband stimulation signal. Our approach did not use a modulated 16 kHz carrier like the House single channel device, but used the broadband analog signal proper for stimulation. Dynamic range compression was achieved by a fast attack/slow release automatic gain control with an adjustable compression ratio. The frequency response was adjusted to closely fit the frequency characteristic of the particular channel/site used. This fitting was achieved by continuously presenting at MCL-level 10-s sweeps over the audio frequency range while simultaneously displaying the frequency response on screen. Thus the patient could on the spot indicate frequencies where adjustments were needed.

This strategy had to cope with the limited-benefit reputation of other single channel devices, but the speech understanding it provided was at least as good as, e.g. the widely promoted F0/F1/F2-strategy. This fact was recognized quite late, following the publication of independent test results by Tyler (1988). Video clips of subjects playing an instrument demonstrate astonishing music perception. This is not surprising in the light of the more recent findings (Smith et al., 2002; Xu and Pfingst, 2003).

Nevertheless, the lack of spectral information limited the achievable speech understanding. Attempts to provide place pitch information in addition to temporal coding by others and by us did not produce the expected improvements. This was either due to the increased channel interaction with simultaneous multichannel analog stimulation (Eddington, 1980), or due to the use of – against our better knowledge – feature extraction to determine channels, i.e. stimulation sites, according to formants F1 and F2 for a pitch-synchronous pulsatile stimulation signal in addition to the analog broadband channel (Zierhofer et al., 1993). A schematic representation is shown in Fig. 2.

The development of the Continuous Interleaved Sampling (CIS) strategy by Blake Wilson and colleagues in the early '90s (Wilson et al., 1991) was a breakthrough. Despite being deceptively simple compared to previous and contemporary feature-extraction strategies, CIS has nevertheless provided impressive improvements in speech perception with cochlear implants. CIS has practically developed into a standard, and the principles behind CIS (frequency analysis, envelope extraction, constant-rate stimulation) have been the foundation of almost every further development in this area. Its success can be at least partially attributed to three features: